REVIEW OF COLUMBIA RIVER TEMPERATURE ASSESSMENT: SIMULATION METHODS

NOVEMBER 1999

Prepared for:

Potlatch Corporation
Lewiston, Idaho

Prepared by:

Peter Shanahan, Ph.D, P.E.
M. Bruce Beck, Ph.D.

481 Great Road, Suite 3
Acton, Massachusetts 01720
(978) 263-1092
fax: (978) 263-8910
e-mail: shanahan@ma.ultranet.com

HydroAnalysis, Inc.

INTRODUCTION

This document provides a review of the report "Columbia River Temperature Assessment: Simulation Methods" prepared by John Yearsley of U.S. Environmental Protection Agency Region 10 in February 1999. The report employs sophisticated modeling and parameter estimation techniques to assess temperature in the Columbia River system and place bounds on the uncertainty in predicting temperature.

The introduction to the EPA report lists the following three sources that may contribute to changes in the temperature of the Columbia and Snake Rivers: impoundments, hydrologic modifications, and watershed modifications. The report does not consider a potentially significant contributor to temperature change, global climate change. The report indicates that it is the objective of the study "to assess the relative importance of these sources with respect to changes in the temperature regime of the main stem Columbia River in Washington and Oregon and in the Snake River in Washington." This objective is partially accomplished. The report limits its focus to the effects of dams and a blanket consideration of contributing tributary temperature.

The number of simulations in the study and their sophistication are insufficient to achieve fully the stated objectives and the study should be viewed as a screening level analysis to identify potential factors that affect river temperature. The report notes on page 5 that the model is indeed a screening model, designed only "to identify critical areas for additional analysis." Thus, the model results should be recognized to be approximate and exploratory rather than definitive. This report would be an inappropriate basis for policy decisions other than the identification of areas for further research.

MATHEMATICAL MODEL DEVELOPMENT

The author elected to write a new computer program for modeling temperature rather than relying upon an established model such as QUAL-2E (Brown and Barnwell, 1987). As explained by author, his approach provides some legitimate advantages for the problem in question. Nonetheless, the failure to use an established model that has experienced substantial use and prior review places an added burden on the author to verify his model.

Typically, computer-model verification is accomplished by using the model to simulate one or more problems for which there is a well-known analytical solution. Close agreement between the model solution and the analytical solution "verifies" the model; that is, it demonstrates that the model accurately solves the type of problem that it was designed to solve (National Research Council, 1989, pp.235-236). The greater the number of such test cases performed, the more confident one can be of the model's accuracy and validity. In essence, verification is a process to guarantee against erroneous computer code. It protects against the situation in which erroneous code begets erroneous results, known in computer science by the shorthand, "garbage in, garbage out."

No verification of the Columbia River temperature model was provided in the reviewed document. Use of the model for screening purposes does not obviate the need for verification. Without such verification, the model's accuracy cannot be assessed and the model cannot be relied upon. The U.S. EPA should be requested to provide documentation of model verification in accordance with accepted standards.

Response: A number of tests of the computer code were performed. In addition, the frequency response of the numerical method was evaluated and compared to the frequency response of other models that have been used for policy-making in the Columbia and Snake rivers. The results of these tests have been added to the Report.

While potential programming errors could compromise the model's validity, the model has a solid technical foundation. The model equations as presented in equations 1, 2, 4, 13-15, and 17-21 are well known and accepted. The equations for stream temperature are solved in the EPA model using a mixed Eulerian-Lagrangian solution scheme. This approach is much less common in surface-water-quality modeling than an Eulerian approach, but is nonetheless entirely valid. The Eulerian approach, which is used for example in QUAL2E, solves for temperature at fixed points along the river. In essence, the Eulerian solution looks at the entire river at a single time. In the Lagrangian approach, the numerical solution follows a particular parcel of water down the river. In essence, this technique looks at this single parcel, following it in time and space as it moves downriver with the streamflow. Travel time for the parcel, τ, and its location along the river, x, are directly related by the stream velocity, U: $x = U\tau$. The

Lagrangian approach takes advantage of this relationship to eliminate the distance variable from the equations being solved.

The mixed solution in the Columbia River model uses a Lagrangian algorithm to model the effect of water flowing downstream in the river, but an Eulerian algorithm to model mixing. The main advantages of the mixed Lagrangian-Eulerian technique, in general, are accurate representation of dispersion and the ability to model sharp transitions in concentration. These are not factors in the current model version since diffusion is not modeled and the concentration (temperature) changes gradually. Thus, although these factors are cited on page 10 of the report, the real advantage of using this solution in the Columbia River model seems to be that it simplifies the accompanying Kalman filtering analysis of model parameters. The role of Kalman filtering in the Columbia River model is further described below.

DATA SOURCES AND PARAMETER ESTIMATION

A multi-step process that considered the statistical properties of temperature prediction and measurement and their uncertainty was used to establish the various parameters in the model. First, model parameters were set based directly on the available data and literature. This first step is called the deterministic parameter estimation. Next, these parameters were adjusted to produce results deemed "unbiased." In essence, this is the same as the manual calibration process used in traditional water-quality modeling. Finally, the variance of the systems model (i.e., the model for river temperature) was estimated as the third and final step of the parameter estimation process. Variance is a statistical measure of the variation of the model results around its mean prediction and depends upon the error in both the model and field data.

The Columbia River model requires a large number of input data to represent the river's geometry and the basin hydrology and meteorology. As discussed on page 13 of the EPA report, numerous past studies of this river system provide an unusually rich database for modeling. The first step in the model development, deterministic parameter estimation, was thus a fairly straightforward and familiar process. A few exceptions are discussed below.

On page 14, the report describes the basis for estimating the river's hydraulic parameters as a function of streamflow under the dam removal scenario. These estimates were made for flows of 60,000, 120,000 and 240,000 cfs in the Snake River. Unfortunately, the minimum flow of 60,000 cfs exceeds the average monthly flow in the Snake River below Ice

Harbor Dam for the months of July through March and is more than twice the monthly average flows for August, September, and October. The model hydraulic parameters thus may be inappropriate for modeling the dam removal scenario during the critical low-flow months. This is significant because these months are also the most likely to experience temperature exceedances.

Response: The hydraulic parameters for the unimpounded Snake River, estimated for the flows of 60,000, 120,000 and 240,000 cfs were used to estimate the hydraulic conditions for 10,000 cfs. These estimates were compared with the results of simulating steady flow with HEC-RAS at flows in the Snake River of 10,000 cfs. A comparison of the two methods showed that the hydraulic parameters used in the report gave results similar to the simulated results from HEC-RAS.

In computing surface heat transfer, the author has appropriately opted to use a heat budget model when simpler, but less accurate, approaches are sometimes used. The heat budget model, which is presented on page 15 and in equations 17 through 21, is based on an authoritative reference (Wunderlich and Gras, 1967) although a somewhat later and more accessible version of the same reference exists (TVA, 1972). The presentation in the EPA report is incomplete, however, in that the report fails to define all variables and units in the equations. This creates some uncertainty as to exactly what relations were actually used. One critical factor in calculating surface heat transfer involves the evaporative heat flux in equation 19. There are multiple formulas for the dependence of this flux on wind speed and the formula chosen may significantly alter the computed heat transfer (Shanahan, 1985). The EPA report fails to clarify the exact formulation used. With respect to the shortwave radiation in equation 17, the atmospheric attenuation coefficient has been shown in some studies in the Pacific Northwest to take on unusual values owing to haze (Findikakis *et al.*, 1980). This may be a factor in some areas of the watershed, which would necessitate special parameter selection.

Response: Peer reviewed methods for determining the various components of the heat budget were used in this report. Additional discussion of parameters, including definition of variables and their units, was added.

On page 15, the report states that **daily** temperature values are not always available for the upstream stations that form the model boundaries. The report then presents, on page 16, an empirical equation for **weekly** temperature that was used to fill in missing temperature values. The mismatch between the daily and weekly periods is significant in that there may be a significant time lag between meteorological conditions and the resulting stream temperatures. As a result, the relation for weekly conditions is likely to be substantially different from the relation for daily conditions. It is unclear from the report whether daily temperature values were actually derived or whether weekly temperatures were used during periods of missing data. In either case, there is substantial margin for error in fixing the temperature at the model inflow points.

Response: A peer reviewed method was used to estimate weekly water temperatures for tributaries with missing data. The report acknowledges the uncertainty associated with these estimates.

Under the subheading "Systems Model Bias and Error" starting on page 17, the report describes how Kalman filtering was used. Kalman filtering augments the supposedly certain prediction made by a traditional water-quality model with a probabilistic prediction that recognizes multiple sources of uncertainty. In a traditional model, the temperature at some time step k is determined from the temperature at the preceding time step k-1, using the model equations and known parameters. In fact, however, the temperature at time step k-1 is known imperfectly, the parameters for stepping from time step k-1 to time step k are uncertain, and even the model formulation itself probably has errors. If the model results are compared with field measurements, the field measurements must also be recognized as having some error. The Kalman filtering approach recognizes these various sources of error and incorporates them into the model formulation. The result is an estimate of the uncertainty in the model predictions that can be used to help guide the calibration procedure.

Kalman filtering is a complicated and specialized technique. Accordingly, expert review of the EPA's application of Kalman filtering in the Columbia River temperature analysis was sought from Professor M. Bruce Beck of the University of Georgia. Dr. Beck is an internationally-known specialist in the application of Kalman filtering to surface-water quality. Dr. Beck's review is appended to this review. Dr. Beck finds no fault in the technical aspects of the Kalman filter analysis, but raises some cautions as to the interpretation of the results. These cautions are pointed out in the discussion that follows.

The EPA study used Kalman filtering in an approach that closely follows that presented by Van Geer *et al.* (1991) for a ground-water modeling application. Despite the change in environmental medium, the approach remains valid for the EPA application. As presented by Van Geer *et al.*, Kalman filtering provides information on the uncertainty of the model prediction and helps guide the calibration process. According to Van Geer *et al.*, one can achieve a better calibration using Kalman filtering than the traditional, deterministic approach. In a persona communication, Dr. Beck has indicated his strong disagreement with this assertion.

Response: The approach used in the report was based on a number of peer reviewed studies, including that of Van Geer et al (1991).

The Kalman filtering procedure is complicated, as is implied in equations 5 through 12 on page 9, and the description in the EPA report is spare and difficult to follow. Accordingly, the procedure is summarized in the next three paragraphs, which are based on the more lucid description in the paper by Van Geer *et al.* (1991). These three paragraphs, which are quite technical, can be skipped without loosing the overall sense of this review. With respect to the procedure described in the EPA report, equations 8 and 12 include errors. In equation 8, the second instance of f_{k-1} should instead be its transpose, $f_{k-1}T$. In equation 12, $\underline{v}_k$ and $\underline{z}_k$ are vectors and should both be underscored.

The Kalman filtering process, as described in equations 5 through 12, marches through time in discrete time steps. It consists of two sub-steps at each time step: first a prediction is made strictly from the model equations, and second it is corrected based on the measured data.

The first sub-step is the prediction. At each time step k in the temperature simulation, the temperature at the various measurement locations along the river (represented by the vector $\underline{T}_k$) is predicted with equation 7 as a function of the system matrix (f_{k-1}) and the temperature at the last time step, $\underline{T}_{k-1}$. The system matrix is simply the temperature equations 2 and 4 in another form. In parallel with the temperature prediction at time step *k*, the uncertainty in the prediction estimated with equation 8. The uncertainty in the temperature is a matrix, P_k, in which the diagonal elements are the variances of each temperature value (i.e., the temperature at each station) and the off-diagonal elements are the covariances between the temperatures at different stations. This matrix is known as the error covariance. Like the temperature vector, the error covariance matrix is predicted based on its value at the last time step.

Following the predictor sub-step is the corrector or update sub-step. Here, the predicted temperature is updated with the actual temperature measurements, $\underline{z}_k$, using equation 9. Equation 9 is simply a weighted average of the predicted temperature and the measured temperature, but with the weighting changing as the simulation progresses. The weighting is captured in the so-called Kalman gain, K_k, which is also a matrix and is calculated in equation 11. In the actual computational sequence, equation 11 is completed before equations 9 and 10. The error covariance is similarly updated in equation 10. At the end of the update sub-step, the calculation for time step *k* is completed and the process begins again with the predictor sub-step for time step *k+1*.

An outcome from Kalman filtering is the innovations sequence, equation 12, which shows the error between the measured and predicted temperature at each location along the river at each time step *k* in the simulation. The goal in calibrating the deterministic temperature model is to adjust the model so as to minimize the mean of the values of this error term over time. In addition, the error covariance matrix (the Σ_Q term in equations 5 and 8) is adjusted until the innovations sequence satisfies certain statistical properties discussed below. Varying, one at a time, the deterministic model and the properties of the stochastic model error adjusts the model. As described in the EPA report, the only deterministic parameter varied was the meteorological data station assigned to each reach of the river. The assignment of stations was varied manually until, according to the report, "the mean of the innovations vector was small." No specific description of "small" is given, although Figures 6 through 13 allow a visual evaluation of the error. Dr. Beck cautions that Figures 6 through 13 appear to compare updated temperature predictions, and thus may present a more favorable comparison to the field data than is appropriate. As indicated in Dr. Beck's review, the exact character of the simulated values in Figures 6 through 13 should be clarified.

Response: Additional discussion of the innovations sequence and the filtering approach was added to clarify these issues.

The stochastic error term that was varied is the estimate of the error in the system model. This error is represented by $\underline{w}_{k-1}$ in equation 5, and it is assumed to be a Gaussian distribution with zero mean and variance Σ_Q. This error is not known at the start of the modeling exercise, so it is given an initial guess and then corrected by trial-and-error based on the results of the Kalman filtering. The corrections entail changing the values of the assumed statistical variance matrix, Σ_Q. The stochastic part of the model is determined to be calibrated when the values of Σ_Q cause the model error computed from the innovations sequence to match the theoretical error predicted by the model. Mechanically, this match is computed using equations 23 and 24 on page 17. As Dr. Beck points out in his review, the values of Σ_Q are expected to differ between simulations of the existing situation with dams and predictions for a future situation without dams. However, it appears that the same values of Σ_Q were used for both scenarios. Dr. Beck also points out that assumptions made concerning the character of covariance terms in Σ_Q are inadequately discussed in the report.

Response: The report acknowledges the uncertainty associated with assuming the systems model error variance, Σ_Q, is the same for all scenarios. This assumption is a result of the absence of water temperature data for scenarios with dams not in place. However, this assumption does not necessarily mean that the variance of the state estimates is the same for all scenarios. The systems dynamics play an important role in the propagation of uncertainty. The systems dynamics for the scenarios with dams in place differ from the dynamics with dams removed and will, in general, give different results for the variance of the state estimates.

The model results raise some questions. The text indicates that data were available for the period 1975 through 1995, but calibration results for only 1990 through 1995 are shown. I cannot be determined from the report whether the entire period of record was used to calibrate the model or if only the 1990-1995 subset of the record was used. It would not be inappropriate to base the calibration on the 1990-1995 period only, since page 16 indicates the data are more reliable then, but the data selection should be clarified.

Response: In the revised report, data collected in association with the total dissolved gas monitoring program for the period 1990-1994 was used in the parameter estimation process. These data were chosen because of their completeness and reliability.

The innovations sequence is a measure of the difference between the temperature predicted by the model and that actually measured. The innovations sequence is shown in Figures 14 through 21 over the calibration period at a number of measurement stations along the rivers. The error is relatively large—greater than 3 or 4 degrees. Moreover, the figures plot a 30-day moving average, implying that some daily values are even more in error. The report is deficient in explaining the meaning, significance, and limitations of these results. The figures illustrate the calibration of the deterministic model where the goal is to get the mean of all plotted values to equal zero. This can be equivalently thought of as getting the area under the plotted curve above the x-axis (0-degree line) equal to the corresponding area below the x-axis. It appears that at some of the stations, the calibration fell well short of this goal. The peaks and valleys in Figures 14 through 21 indicate that the model appears consistently to predict temperatures that are too low in fall, winter, and spring, but too high in summer. Dr. Beck furthe discusses the lack of coherence between Figures 14 through 21 and Figures 6 through 13 and its implications insofar as relying upon the model predictions.

Response: A number of statistics relating to model performance have been added to the report.

The report is similarly deficient in explaining the meaning, significance, and limitations of the results in Figures 22-29. In essence, these figures report on the calibration of the stochastic model, plotting the results of equation 23 against those of equation 24. A goal of model calibration is to get these to match. Unfortunately, the key of these figures is insufficiently clear to distinguish which plotted line represents which result. The terminology of the figures deviates from that of the text, further confusing the results. As with the results in Figures 14-21, the results in Figures 22-29 show significant variations over time and, in at least some cases, a consistent mismatch.

Based on the comparisons in the figures, it is difficult to assess the quality of the calibration. More information on alternative calibration attempts would be helpful in this regard and would also give a sense of the model sensitivity. As well, segregation of the model-data comparison by month would help in identifying the accuracy with which the highest temperatures are predicted. Model predictions are particularly critical in this range because it is only this portion of the model results that are actually evaluated.

MODEL APPLICATION

On page 18, the report states goals that are not entirely congruent with the objectives stated on page 1. Also incongruent are the conclusions on page 20. While page 5 states that the model is a screening tool capable of identifying areas for further study, the report make no recommendations for further study. Instead, the report lists seemingly firm conclusions—an outcome that is inconsistent with the power and purpose of a screening model.

Model results are shown in terms of the frequency with which a temperature of 20°C is exceeded at the various stations along the rivers (in Figures 30-35) as well as the degree to which the temperatures are exceeded (in Figures 36-41). Simulation scenarios consider the current situation, the situation if existing dams were to be removed, and the situation if temperatures from tributary streams were kept less than or equal to 16°C. The simulations show that the frequency and magnitude with which 20°C is exceeded is decreased by removing dams (other than at the Snake River confluence and Grand Coulee Dam) and relatively unaltered by controlling tributary temperature.

There is a significant mismatch between the way the model was used and the way it was developed that calls into question its predictions. In its use, the model is applied only to evaluate extreme high temperatures that occur in the summer. But, the model's calibration and statistica evaluation were judged in terms of year-round agreement. The statistical measures used in the Kalman filtering evaluate the degree of agreement over the entire year and, for the deterministic model, via summations over the entire calibration period. Thus the summertime predictions, which tend to be high, are offset by the non-summer predictions, which tend to be low. The results presented in the report, however, show extreme temperature exceedances that occur only in the summertime period. Before the model can be confidently used to evaluate temperature extremes, it must be calibrated and checked specifically against periods of high temperature. Dr. Beck confirms this conclusion in his review.

This fundamental limitation notwithstanding, the model results predict temperature exceedances (in Figures 36 through 41) that are comparable to the calibration errors depicted in Figures 6 through 21. The "error bars" shown in Figures 36 through 41 may be confusing in this regard. They show the variation of the predicted exceedances around the mean and do not relate to the model uncertainty. However, it is clear from inspection of Figures 6 through 13 that the temperature model makes its poorest predictions at the extremes, yet it is precisely at the extremes where the model is being used.

SUMMARY

The EPA Columbia River temperature model uses unusual and technically sophisticated techniques to evaluate the effects of dams and other factors on temperature in the Columbia and Snake Rivers. Because an established model was not used, the Columbia River model should be verified in accordance with accepted practices for model quality assurance and quality control.

Calibration information provided for the model appears to show that the model predicts summertime temperatures that are generally higher than those observed and non-summer temperatures that tend to be lower than observed. However, the model calibration was evaluated in terms of year-round agreement, such that these two systematic errors balance each other. In contrast to the calibration evaluation, the model was used in a predictive mode only to evaluate extreme warm temperatures in the summertime. If the model is to be used primarily or solely to evaluate high temperature extremes, its predictive capability should be evaluated specifically for high temperature.

Errors in the model during summer appear to be comparable to the degree of exceedance predicted for summertime temperature excursion above the 20°C temperature threshold. This relative similarity of model error to the predicted excursion, as well as the mismatch between the calibration focus and prediction focus, indicate that the model results should be considered qualitative at best. As indicated in the report itself, the model is intended as a screening tool to identify areas for further research. As such, the model is not an adequate basis for policy decisions.

In a separate appended review, Dr. M. Bruce Beck focuses on the application of Kalman filtering in the EPA study. Dr. Beck concludes the Kalman filtering is implemented in a technically sound manner overall, but that certain aspects of the application require clarification He also questions a number of explicit and implicit assumptions regarding the character of error and uncertainty and suggests additional analysis to explore their consequences.

REFERENCES

Brown, L.C., and T.O. Barnwell, 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual. Report No. EPA/600/3-87/007. U.S. Environmental Protection Agency, Athens, Georgia. May 1987.

Findikakis, A., F.A. Locher, and P.J. Ryan, 1980. Temperature and Turbidity Simulation in Spada Lake. In: H.G. Stefan, editor. *Proceedings of the Symposium on Surface Water Impoundments. June 2-5, 1980, Minneapolis, Minnesota.* Vol. I, Pg. 594-603. American Society of Civil Engineers, New York.

National Research Council, 1989. *Ground Water Models: Scientific and Regulatory Applications.* National Academy Press, Washington, D.C.

Shanahan, P., 1985. Water Temperature Modeling: A Practical Guide. in T.O. Barnwell, Jr., editor. Proceedings of the Stormwater and Water Quality Model Users Group Meeting, April 12-13, 1984. Report EPA-600/9-85-003. U.S. Environmental Protection Agency, Athens, Georgia.

TVA, 1972. Heat and Mass Transfer Between a Water Surface and the Atmosphere. Water Resources Research Laboratory No. 14. Tennessee Valley Authority, Division of Water Control Planning, Engineering Laboratory, Norris, Tennessee. April 1972.

Van Geer, F.C., C.B.M. Te Stroet, and Y. Zhou, 1991. Using Kalman Filtering to Improve and Quantify the Uncertainty of Numerical Groundwater Simulations, 1. The Role of System Noise and Its Calibration. *Water Resources Research*, Vol. 27, No. 8, Pgs. 1987-1994. August 1991.
J232-rep1.doc

Wunderlich, W.O., and R. Gras, 1967. Heat and Mass Transfer Between a Water Surface and the Atmosphere. Tennessee Valley Authority, Division of Water Control Planning, Engineering Laboratory, Norris, Tennessee.

APPENDIX

Review of Kalman filtering
by Dr. M. Bruce Beck